



*Citation for published version:*

Daniels, J 2017, 'A matter of opinion: the Delphi method in the social sciences' *SAGE Research Methods Cases*.  
<https://doi.org/10.4135/9781526410603>

*DOI:*

[10.4135/9781526410603](https://doi.org/10.4135/9781526410603)

*Publication date:*

2017

*Document Version*

Peer reviewed version

[Link to publication](#)

## University of Bath

### Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# SAGE Research Methods Case Health Submission for Consideration

---

## **Case Title**

A Matter of Opinion: The Delphi Method in the Social Sciences

## **Author Name(s)**

Dr Jo Daniels

BSc (hons) ClinPsyD CPsychol BABCP FHEA

## **Author Affiliation & Country of Affiliation**

University of Bath, UK

## **Lead Author Email Address**

Email: j.daniels@bath.ac.uk

---

**Discipline: D4 [please do not alter]**

## **Contributor Biographies**

Dr Jo Daniels is a Chartered Clinical Psychologist who qualified from the University of Manchester clinical psychology training programme in 2009. Since qualification she has worked across health and mental health settings before gaining her accreditation as a Cognitive Behavioral Psychotherapist in 2011 and moving to the University of Bath Clinical Psychology Training Programme as a lecturer and tutor. She currently teaches as a Fellow of the Higher Education Academy on both postgraduate and undergraduate psychology courses. Dr Daniels is research active in the field of clinical health psychology, with a special interest in anxiety and cognitive behavioral models.

## **Associated Published Article**

Roos, J. & Wearden, A. (2009) What do we mean by 'Socialization to the model': a Delphi study, *Behavioural and Cognitive Psychotherapy Journal*, 37 (3), 341-345

## **Abstract**

In 2006, as part of my clinical psychology doctoral thesis I sought to examine predictors of outcome in psychological treatment, however I was unable to identify an operationalized definition of a key term from which I needed to base my research. “Socialization to the model” was a phrase used ubiquitously in clinical textbooks however the definition differed and was too nebulous to test. It was essential to the success of my research that I adopt a rigorous approach and ask the question: what do people mean when they use the term ‘socialization to the model’ and more importantly, do they mean the same thing? The answers to these questions would allow me to test an essential ingredient in the therapeutic process.

In this case study I describe and critically appraise a methodology which allowed me to systematically obtain opinions of experts to produce a robust and statistically reliable definition of the term which was set to underpin my future research: the Delphi method.

I discuss the advantages of the mixed methods approach which offers flexibility from a number of perspectives, an advantage which deserves a cautious appreciation. I examine the inherent challenges of the approach, specifically the avenues for bias, and highlight the essential role of the sample, drawing on my own reflections and experiences.

## **Learning Outcomes**

By the end of the case, students should:

- Understand and describe the Delphi methodology and how this can be applied in the field of social sciences
- Understanding the inherent methodological advantages and challenges within the Delphi approach as applied to the social sciences and more broadly
- Be able to critically appreciate the avenues of bias in the Delphi approach and be able to consider methods to control for these biases

- Be able to critically appraise this mixed methods approach in comparison to other qualitative and quantitative approaches

### **Case study**

As part of my clinical psychology doctoral training I conducted a research study examining predictors of psychological treatment outcomes in Chronic Fatigue Syndrome (CFS), however during the initial stages of the project I found myself at a loss; I was passionately invested in my ideas about what might be important to treatment outcomes yet I was struggling to define the key concepts. The term ‘Socialization’ is commonly used by psychological therapists to describe a process whereby a patient and therapist arrive at a shared understanding of a presenting problem and an agreement around what they have to do to deal with that problem. The reason I know what that term means, is because I had to define it.

Although the term ‘socialization’ was clinically and academically ubiquitous, it was ill-defined and certainly not operationalized enough to test my theory that ‘socialization’ was a predictor of treatment outcome in CFS. It was clear to me as a clinician that having an accurate shared understanding of a patients’ presenting difficulties was essential to therapy, whether it be anxiety, depression or something else; how could you and the patient work effectively together if you have different ideas about what the problem is? It would be like going to the doctor with a broken arm, and coming out with a bandaged leg – the treatment is dependent on the clinician understanding what the problem is and the patient effectively communicating the symptoms. It also felt important that the patient understood the implications of the treatment, so the patient knew what they were ‘buying’ into.

However, the discrepancies between definitions and a lack of research in the area gave rise to a need to operationalize the term for my research purposes.

Wells (1997) referred to socialization as ‘selling the cognitive model and providing a basic mental set for understanding the nature of treatment’. However, Wells did not provide a definition that could be utilised in research or clinical practice.

In psychological therapy research methodologies can be broadly divided into qualitative and quantitative approaches, with some exceptions which combine little of one and more of the other. As this area was an under-researched there was no validated questionnaire available to use and no empirical evidence on which to develop a questionnaire. Qualitative interviews would be a labour intensive task for such small scale deliverables; grounded theory and thematic analysis would of course yield very rich and complex data, however it didn’t quite fit with my need. I required a methodology which enabled me to operationalize, or define, if you will, a commonly used term so that I could robustly test my research hypothesis. I was in a quandary.

Reading around the area of psychological therapy, I stumbled upon a paper by Duncan and colleagues (2004) which used the Delphi method to elicit opinions and consensus agreement over what should be covered in a psychological therapy manual. I was immediately engaged with the very versatile and innovative methodology which was previously unknown to me: the Delphi.

An approach initially developed by Dalkey and Helmer (1963) at Rand Corporation for business and economic forecasting, this approach took some time to penetrate mainstream

research methodology and broaden its application to other fields. However, there are now many papers written on the approach, and social scientists appear to include this unique technique in their methodology tool kit. Indeed, it has been used by the National Institute for Clinical health and Excellence (NICE, 2004) in the development of guidelines for clinical interventions in the UK National Health Service.

The approach itself takes the form of a multi-stage iterative group communication designed to elicit consensus agreement on a given subject. In this case, I needed experts in the field of psychological therapy who could tell me what they meant by ‘socialization to a treatment model’. If consensus was achieved, I would then be able to offer an evidence-based definition for clinical application, and also proceed to test my research hypotheses about its relevance to treatment outcomes

## **Research design**

The Delphi usually consists of three rounds, each round building on the previous with the objective of allowing experts to anonymously refine their views in light of views expressed by other participants, as the group’s work progresses (Millar et al. 2006). In other words, as the participants express their own opinions, and are exposed to others opinions, there may be convergence and confluence of ideas as the rounds progress. The aim is to reach a consensus agreement on a specific topic or question through this iterative process. Although the three round Delphi is most commonly adopted, rounds may continue if a higher level of consensus is required.

The panel of 'experts' are the participants and sample within this method. Experts can be systematically identified by using clear inclusion criteria, for example they must have published and be currently working in the associated field, or it may be that you want to develop a consensus agreement on a topic with a specific target audience, such as a team or organization, to gain agreement on a specific topic. The size of the panel of experts is also a pertinent consideration. Rowe and Wright (2001) suggest that between 2-12 participants is sufficient for a Delphi study, however this is relative to your research question and your target audience.

The procedure of the Delphi itself is most commonly delivered via postal, or email correspondence. While email may be more time-efficient approach, the disadvantage is compromised anonymity to the researcher. This may not be an issue as experts may be confident in their opinions, however it may be a worthy consideration for sensitive matters.

The initial round of the Delphi takes the form of a questionnaire based on the research question(s), with an accompanying free text response to these questions. This is designed to elicit individual responses to the questions.

At this stage, all responses are collated and collapsed into statements. For example, there may be 12 participant responses, however due to the overlap it is possible to collapse the qualitative data from the free text responses down to 7 statements. It could be suggested that the collapsing of common statements is subject to bias however one of the advantages of this approach is that the second stage allows all participants to edit or modify the statements at each round, which controls for bias in interpretation to some degree.

The second round moves more into the quantitative domain; a questionnaire is devised based on the common statements and participants are invited to edit/modify the wording or phrasing, but also indicate their level of agreement with the relevance and importance of the statement. The question you are asking is ‘how important are these common statements (that you and your peers have generated) in answering this question?’ This is usually done using a Likert scale (e.g. 1-5) where 1 might represent ‘irrelevant’ and 5 might represent ‘essential’. The rating and wording would of course depend on the question you are asking. Analysis of round two involves the integration of suggested modifications and additions of any new statements, and a calculation of the median scores of each statement.

The final round, round three, is based on the newly modified statements. Participants are invited to make further edits or modifications and then rate the statements for a second time, stating their reasons for any rating that was more than one point different from the round two median rating.

The analysis is then divided into two areas: level of agreement and level of importance given to the statements.

Median and inter-quartile ranges are calculated for each statement in round 3 responses.

Statements with an Inter-Quartile Range (IQR) of  $<1$  can be defined as having reached consensus, and an IQR of 0 indicating absolute consensus. So this gives an indication as to how much the group agrees with one another. If the interquartile range is wide, then this reflects that the group members do not agree on a specific statement. When the range is  $<1$  or 0, you have an indication that experts share an understanding and agreement on a given



statement. This will result in a smaller pool of statements which collectively reflect agreement around a given subject.

The second aspect of analysis is the examination of what the participants actually agree on. The final statements with affirmed consensus may all agree that the statements are essential or important, or they may universally agree that a statement is irrelevant to the research question, or somewhere in between. These are the possibilities, however the likelihood of poor agreement is reduced due to the participants themselves generating and evolving the statements over the course of the rounds. This is one of the advantages of the Delphi over traditional questionnaires; the participants themselves produce the data which they then later rate in a questionnaire format, however you still emerge with a quantified statistically reliable outcome. In fact, Fowles (1978) identifies anonymity, controlled feedback, and a statistically valid response as the main characteristics and advantages of the Delphi technique.

The final stage is the report on the culmination of the Delphi rounds. The shape of the outcome or product is again dependent on the nature of the research question, and may take the form of a list of agreed items for inclusion in a therapy manual, a definition of a concept, a description of the shared goals and objectives of a team, or something else. This is the real beauty of this approach; the versatility and flexibility.

There are many publications around the use of the Delphi; I would recommend the concise summary paper by Hsu and Sandford (2007) and the fuller detailed text by Linstone and Turoff (2002).

### **Method in action**

My research question was clear and simple from the outset. I wanted to know what people meant when they used the term ‘socialization’ when referring to an aspect of psychological therapy, and whether they meant the same thing. The Delphi approach was considered to be a highly appropriate method to use to answer this research question given the paucity of research in the area, so I embarked on my first, but not my last exploration of meaning using the Delphi technique.

### *Sample*

Based on previous guidelines and recommendations around sample sizes, our aim was to recruit 6-8 experts, with recruitment continuing until this number was reached. The inclusion criteria was discussed and agreed by myself and a colleague; as the criteria for ‘expert’ is somewhat subjective and relative, it felt important that more than one person was involved in the development of the inclusion criteria. It was agreed that participants should be qualified Clinical Psychologists with at least two years post-doctoral clinical experience, as this would indicate they were sufficiently qualified and experienced in the relevant field. Criteria two referred to academic prowess; they needed to be either conducting research or teaching/lecturing at doctoral level in the field. This gave an indication that they were likely to have advanced knowledge and expertise in the field and demonstrate a more rigorous understanding of theoretical concepts. This is a potentially controversial claim, however appointment to an academic role was considered important and the most likely place to find experts with up-to-date knowledge. The final criterion was a history of publications (i.e. one or more) on relevant clinical theory or interventions.

Through a review of the literature, sixteen potential participants were identified and contacted via email. Of these, nine agreed to participate by responding to the first round by the two week deadline. It is often found that people respond fairly immediately if they are going to respond at all (Lemon, 2007), hence the two week deadline.

All nine participants completed all three rounds, demonstrating 100% response rate. In terms of sample characteristics, all indicated their approach was broadly Cognitive Therapy based, however one participant identified more strongly with a ‘social constructionist’ approach. Three out of the nine participants were female. No other demographic information was gathered as this was not deemed particularly relevant to the study.

### ***Procedure***

Email correspondence was used to facilitate the Delphi, as this was considered to be convenient for the panel and more likely to elicit a rapid response. In the first round, participants were emailed and asked to respond to the two research questions:

- (1) What is socialization to the model?
- (2) How do you know when a service-user is socialized to the model?

The term ‘service-user’ in this situation refers to a person who is engaged with a therapy process. Others may use the term service-user and patient interchangeably, however at the time of the study this was the commonly accepted term of reference.

### ***Analysis***

The content of the nine email free text responses were analysed independently by me and a colleague. Despite the Delphi consisting of two predetermined research questions in the first round, three sections emerged overall: direct responses to question 1, direct responses to question 2, and additional information regarding how an expert would ‘socialize’ to the model.

This final section was then treated as section 3 of 3, with the Delphi rounds adapted accordingly. This expansion was to encompass any relevant material generated by the panel to illustrate their responses, i.e. their examples. The research was designed to be fairly focussed on the main research question; however it seemed foolish to ignore an emerging theme that could potentially enhance understanding.

For question 1, eleven statements deemed to be in common to all nine responses were extracted and collapsed into four statements representing the components of “socialization to the model”. Below is an example of a raw data free text response:

*Socialisation to the model refers to enabling the patient to understand how their presenting problem is maintained, and the factors that treatment will focus on. This usually involves deriving a **case conceptualization** based on an evidence-based model in cognitive therapy and using verbal and behavioural socialization techniques to illustrate the role of cognition and behaviour in problem maintenance.*

For illustrative purposes, I have included the summary data table from round 1 (see table 1). I have emboldened key aspects of the extract, and the corresponding data on the summary table to highlight the process of collapsing down into common statements.

As we did not specify the treatment model in the question and we felt there may be some generalizability of the term, we kept the definition focused on the term ‘socialization’ which could form part of other therapeutic approaches. This led to the exclusion of ‘evidence based’, ‘cognitive therapy’ or any reference to a specific therapeutic model. We did however extract references to suggested techniques (underlined> which was included in (new) section 3: ‘what techniques would you use to socialize a service-user to the model’? As you can see, the information underlined appears to be valuable and important and too rich to ignore.

All components of the free text response are featured in the raw data column and also in the collapsed statements. As you can see in the raw data column, there is a high degree of overlap, hence the substantially reduced number of collapsed statements.

**Table 1: Raw data and collapsed data from question 1 “What is socialization to the model”**

|                             |  |                                    |  |
|-----------------------------|--|------------------------------------|--|
| <b>RAW<br/>DATA:<br/>11</b> | <b>Understand problem and maintenance in relation to model</b> | <b>COLLAPSED<br/>STATEMENTS: 4</b> | <b>Case conceptualization</b>            |
|                             | <b>Case conceptualisation based on model</b>                   |                                    | <b>Shared understanding of problem</b>   |
|                             | <b>Understanding of focus of treatment</b>                     |                                    | <b>Educate service user /teach model</b> |
|                             | Case conceptualisation   |                                    | <b>understands focus of treatment</b>    |
|                             | Shared understanding of objective of treatment                 |                                    |  |
|                             | Shared understanding of how objective will be met              |                                    |  |
|                             | Explain model so SU  |                                    |  |

|  |   |  |
|--|---|--|
|  | understands                                 |  |
|  | Explanation of difficulties in formulation  |  |
|  | SU understands expectations of therapy      |  |
|  | Shared guided understanding of difficulties |  |
|  | SU is educated about model and principles   |  |
|  |   |  |

For question 2, twenty-seven statements were extracted from the nine responses and collapsed into eleven statements representing indicators of “socialization to the model”. For section 3, nine statements were collapsed into six statements.

The twenty-one statements derived from the original question were divided into the corresponding three sections and used to construct the second round Delphi questionnaire in which participants were asked to rate each statement on a five point Likert scale, ranging from “irrelevant”=1 to “essential”=5. This same scale was used in response rounds, 2 and 3.

Based on feedback from a participant in round 2, an additional rating of “Not applicable” was added to the rating system.

After round 2 had been completed, one statement was added in response to question 1. No other statements were added, although wording was edited by two participants. As planned,

median scores on each of the twenty-one statements were calculated following full sample completion of round 2.

In round three of the Delphi study, participants were presented with the sample median score for each of the twenty-one statements. They were then asked to rate the statements for a second time, and to state their reasons for giving a rating of more than one point difference from the round 2 median rating, if this was applicable. The responses from this round completed the Delphi study data collection.

The median and inter-quartile range for each statement was calculated for round 3 responses.

## ***Results***

Criteria used for the development of a working definition was inclusions of all statements rated as “essential” and “important” with an IQR  $<1$ . For question 1, this included all five statements derived from the round 3 Delphi. Five out of eleven indicators relating to question 2 of the Delphi questionnaire, were rated as “important” indicators of socialization and also had IQR of  $<1$ . Statements 1 and 2 had an IQR of 0 indicating complete consensus. The statements were adapted to form a comprehensive definition to retain all relevant information:

*‘Socialization to the model’ is the process by which a service-user and clinician negotiate a shared understanding of the presenting difficulty. During the process, the clinician presents hypotheses and a formulation of the service-users’ symptoms and experience in terms of the model to be used for the intervention. The therapist provides information concerning the practical implications of the chosen model of*

*therapeutic intervention, to allow the service-user to fully engage with and understand both the therapeutic process and the rationale for intervention.*

Only one statement was rated as 5=essential to understanding how to socialize a service-user to the model of intervention. The IQR for this statement was 0, indicating complete consensus:

*Use techniques and examples to illustrate and provide support for chosen model of intervention/treatment*

One statement was rated as 4=important with an IQR of 1 which indicated good consensus:

*Give information about the expectations of the intervention, including course and outcome.*

Statements were withdrawn if the sample response indicated either a low level of consensus or a high level of consensus that a statement was not important to questions 1,2 or 3.

## **Practical Lessons Learned**

- **Defining your ‘experts’** is a subjective and important process. The Oxford English Dictionary defines an expert as: a person who is very knowledgeable about or skilful in a particular area (OED, 2009). This definition leaves a lot of room for flexibility in terms of what constitutes ‘very’ knowledgeable. Rigour at the sampling stage is highly important and should be considered a risk of bias to be controlled for. While the



outcome may or may not vary dependent on the level of 'knowledge' it is plain good science to be as systematic and robust as you can, even when the parameters are malleable.

- **Work with others:** there is room for bias in interpretation. Often in qualitative studies a caveat is offered by the researcher that saturation in data and reported emergent themes will inevitably be influenced by the researchers own experiences and expectations. The process of collapsing statements down to more concise statements without repetition will involve a level interpretation. Involving a co-rater increases the robustness of the project and also ensures that you are not unduly influencing the content of further rounds. Co-rating is a common control for bias in qualitative in methods, and should be applied when interpretation is required.
- **Respecting the stakeholders:** when I set out to define 'socialization' my main aim was to achieve a well-defined scientifically valid definition for clinical and research use. What I did not anticipate was that by inviting experts to participate, I would be engaging people who were invested in the project and for whom this was important. They were not just participating because they are doing their 'bit for science', it actually mattered to them. This is where duty of dissemination comes in. We are all encouraged to disseminate our research for the greater good of science, however it is also important to remember that to file away your findings is a great disservice to the project stakeholders and scientific community who may be pondering the meaning of nebulous issues you have already explored.

## Discussion

The definition produced through the Delphi method had a high degree of face validity and was consistent with the literature reviewed, which is what we would expect from a panel of experts using a collaborative, iterative process. The findings were drawn together to produce a working definition of the term ‘socialization’ which was potentially useful for clinicians and sufficiently operationalized for my future research. It was evident from the repeated complete response rate and rigour of participants input that the sample were highly engaged in the process, offering editions, corrections and expanded responses to illustrate their opinions. This is the benefit of using ‘experts’ in a field who already have fully formed and highly considered opinions. However, there is a limitation to using experts in the pursuit of quantifying knowledge: the definition that emerges is likely to be a convergence and confluence of opinions in an exclusive sphere; do we have something to learn from non-experts? Could this expand our understanding further? Would there in fact be any differences? A limitation of the Delphi technique is that the product of the study (in this case the definition of socialization) can only reflect the participants who were approached and participated - other “experts” may have influenced the development of the definition in a different way.

A common criticism of the Delphi method is the process by which participants are selected for inclusion: participant selection method is one of the significant components of a reliable and valid Delphi study. In retrospect, our choice of experts was less rigorous than would have been ideal. Myself and a colleague nominate the ‘experts’ based on discussions of influential clinicians and academics in the field. It is likely that this selection method would have been influenced by shared exposure to academics and clinicians through the clinical psychology training and perhaps somewhat bias the choice of nominated experts. I would most likely choose different experts today, however the definition my participants produced

still rings true almost a decade later and I would hedge my bets that the outcome really wouldn't differ significantly.

Where questionnaires are hampered by demand characteristics, the Delphi approach allows the capture of complex data without the inherent difficulties commonly associated with quantitative approaches, yet also produce statistically reliable outcomes. It combines some of the best aspects of qualitative and quantitative methods. The flexibility of the approach permitted a natural evolution over the iterative stages, facilitating a seamless transition from free text initial responses to a highly defined and concise definition of a term in a relatively short period, making the approach easy to administer and manage.

While my sample of experts was relatively small, this isn't a major issue for a study of this nature; the use of a small sample with the Delphi method is appropriate when the study aim is to generate new information on a topic that is generally understood but not specifically defined (Adler & Ziglio, 1996). In this way, it is similar to qualitative interviews, where a small number of participants are interviewed and the emergent themes across a relatively small sample.

The Delphi method is a little known method that is versatile, flexible, efficient and a quantifier of qualitative data. Developing a working definition through the iteration of expert opinion was a revealing and curious process; I was enthralled to see a consensus emerge early on and waited with baited breath to see how the hive mind evolved. There is however room for error and bias which is fundamental to consider in its application, which was not clear to me at the outset, but is evident with reflection.

Not without limitations or criticisms, as no methods are, the Delphi method has carved out a unique place in the social scientists methodology tool kit as a statistically reliable approach which quantifies qualitative data, seeking to gain a consensus on topics that matter. Well, that's just my opinion, other "experts" may think differently!

## Exercises and Discussion Questions

- What other methods could be used to define a commonly used term?
- How can sampling bias be minimized in a Delphi study?
- Why might flexibility *not* be a good thing in research methodology?
- What are the limitations of using small samples vs larger samples for an approach such as the Delphi?
- What are the benefits and drawbacks of using an anonymous approach vs. non-anonymised approach in the Delphi method?

## Further Readings

Linstone, H. A., & Turoff, M. (2002). The Delphi Method. *Techniques and applications*, 53.

Hsu, C. C., & Sanford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical assessment, research & evaluation*, 12(10), 1-8.

## References

Adler, M., & Ziglio, E. (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Jessica Kingsley Publishers.

- Duncan, E. A. S., Nicol, M., & Ager, A. (2004). Factors that constitute a good cognitive behavioural treatment manual: A Delphi study. *Behavioural and Cognitive Psychotherapy*, 32(02), 199-213.
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 3, 458-467.
- Fowles, J. (1978). Handbook of futures research. Greenwood Press: Connecticut.
- Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical assessment, research & evaluation*, 12(10), 1-8.
- Lemon, J. S. (2007). The effect of reminder intervals on response rates for web surveys. *Association for Survey Computing*, 103.
- Linstone, H. A., & Turoff, M. (2002). The Delphi Method. *Techniques and applications*, 53.
- Millar, K., Tomkins, S., Thorstensen, E., Mephram, B., & Kaiser, M. (2006). Ethical Delphi.
- NICE. (2004). Guideline development methods: Making group decisions and reaching consensus [Electronic Version]. Retrieved 15th October 2008.
- Oxford English Dictionary (2009) Oxford University Press: Oxford
- Rowe, G., & Wright, G. (2001). Expert opinions in forecasting: Role of the Delphi technique. In J. Armstrong (Ed.), *Principles of forecasting: A handbook of researchers and practitioners*. Boston: Kluwer Academic Publishers.
- Wells, A. (1997). *Cognitive therapy of anxiety disorders: A practice manual and conceptual guide*. Chichester: John Wiley & Sons Ltd.